DOCUMENT RESUME

ED 198 147 TM 810 061

AUTHOR Williams, Richard C; Bank, Adrianne
TITLE School District Linking of Evaluation and Testing with Instruction: A Problem Worthy of Attention.
INSTITUTION California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE 21 Nov 80
NOTE 17p.: Paper presented at the Annual Meeting of the Evaluation Research Society (Arlington, VA, November 21, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational Testing; Elementary Secondary Education; *Evaluation; *Information Utilization; Instruction; *Management Systems; *School Districts; Systems Development; *Technology Transfer

ABSTRACT

Various notions about the obstacles to, and the conditions needed for, school districts' linking of externally mandated testing and evaluation with instruction are discussed. Two major points are made, the first being that evaluation and testing communities must attend to the needs of administrators and teachers as they are manifest in the district environment, if testing and evaluation efforts are to be used for instructional improvement at the local level. The second point refers to the actual linking of testing and evaluation with instruction: a variety of coordinating mechanisms and a range of related operations are needed to compose district-wide management subsystems. Also highlighted are background information on the growth of evaluation and testing, observations on school districts as organizations, and tentative propositions about school districts' linking of evaluation and testing with instruction. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED198147

SCHOOL DISTRICT LINKING OF EVALUATION AND TESTING WITH INSTRUCTION: A PROBLEM WORTHY OF ATTENTION*

Richard C. Williams and Adrianne Bank

Introduction

Since its inception in 1972, the National Institute of Education has supported research on educational evaluation. This research has contributed to our understanding of how evaluation findngs can improve practice in American schools. As a result of this research effort, it has become clear that the transformation of evaluation findings into appropriate action does not occur automatically. Instead, it is a complex process which is influenced by the individuals who are expected to take action and on the organizational settings within which they work.

Two large groups are possible users of evaluation findings. One group is policy makers such as federal and state legislators or administrators. Evaluators expect that large-scale evaluations of funded educational programs can give this group sound information on which to change program requirements, to augment or cancel programs. A second group includes district and school officials and classroom teachers. Evaluators expect that the data collected in school classrooms can be of benefit to district administrators, teachers, and students.

---

*Presented at the Evaluation Research Society meeting, Arlington, Virginia, November 21, 1980. The project reported here was performed persuant to a grant from the National Institute of Education to the Center for the Study of Evaluation. However, the opinions expressed herein do not necessarily reflect the position or policy of the NIE and no official endorsement should be inferred.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
R. C. Williams
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 810 061

At first glance, it would seem that findings from a given program evaluation would be equally useful to both groups of people. Such appears not to be the case. Studies by Alkin et al. (1979), David (1978), Patton (1978), and Weiss (1972) have found that users' interests and settings influence the reception given to evaluation findings. One major implication of these studies is that evaluators who wish their findings to be utilized must attend in advance to such interests and settings.

The Center for the Study of Evaluation, for the past five years, has been conducting research within school districts. One of our purposes is to describe these organizational settings and those who work there so as to better understand whether and how evaluation and testing are affecting educational practice at the local level. This paper will present background information on the growth of evaluation and testing; some observations on school districts as organizations, our research methods, and some tentative propositions about school districts' linking of evaluation and testing with instruction.

Background

The last 15 years have seen the growth of what might be called a testing and evaluation movement within American education. The seeds for this development were planted decades ago when psychological, intelligence, and aptitude tests to screen and sort individuals were developed by the military and industry and subsequently by school districts. Many large school districts developed test bureaus which regularly collected and disseminated district-wide test results.

These somewhat dormant test seedlings experienced an enormous growth spurt when Congress passed the Elementary and Secondary Education

Act in 1965 requiring systematic program evaluation as a condition of funding. Subsequent federal and state legislation carried similar conditions. Instead of tests being used primarily to assess individual students, they were now to be used to assess groups of students. Instead of test scores being used, for example, to assign individuals to special educational programs, scores were to be used to certify, modify or eliminate the programs themselves. In short, large-scale testing of students had become a tool for program evaluations generating data with which policy makers were to identify, further develop, and disseminate promising educational programs and practices.

The large-scale infusion of federal funds into education since 1965 has had many reverberations. One important side effect of these funds is the flowering of what might be called a testing and evaluation "establishment" both outside as well as inside school districts. Outside school districts this evaluation "establishment" includes: units within federal and state governments devoted to program evaluation and testing; university courses and degree sequences in testing and evaluation; professional societies within education such as Division H of AERA, and across social action fields such as the Evaluation Research Society; a federally funded Center for the Study of Evaluation. Inside school districts the educational evaluation establishment consists of testing and evaluation personnel who face the task of carrying out these newly required evaluation and testing efforts,. Evaluation units have emerged, often composed of already employed guidance, testing, and counseling personnel. Within the last 10 years, over 400 districts have centralized their testing and evaluation capabilities into research, development, and evaluation (RD&E) units which vary in size

from one part-time person to dozens of professional employees (Lyon, 1978).

At first, some members of the evaluation establishment outside of the school districts concentrated their attention on the logic and methodology of large-scale evaluations. Their assumption, not always explicit, was that policy makers would be the prime consumers of evaluation information; school district evaluators had only to collect the data meticulously and accurately and file reports.

However, concern with utilization at the policy level has developed as evaluators have begun to evaluate their own work, seeking evidence that their reports have had the desired policy-making and programmatic impact. It has become increasingly clear that evaluation and testing reports, as they are presently constituted, do not uniformly influence policy makers. Some reports have influence on some policy makers under some circumstances. At other times, the reports are used selectively to provide corroborating evidence for policy makers to justify decisions that they had already made on some other basis. And in still other instances, reports are ignored.

Concurrently, some outside evaluators as well as those within school districts have begun to look to examine whether evaluation and testing data can be not only as a means of satisfying the evaluation and testing concerns of external funding agencies, but also a basis for systematic and comprehensive local school district decision making.

The Center for the Study of Evaluation, established by the same 1965 act which attached evaluation requirements to federally funded programs, was initially oriented largely toward methodological and technical issues. In recent years, projects have been added which are

concerned with evaluation utilization at the program administrator level (Alkin et al., 1979; Daillak, 1980), and at the school district organizational level (Lyon et al., 1978). From these studies of how evaluators relate to clients and how research, development, and evaluation units handle their activities, it has become clear to us that although the potentiality exists, local use of evaluation and testing does not occur routinely as a natural consequence of conducting an evaluation. Some combination of special circumstances, special people, and special organizational arrangements are required to link data collection, evaluation reporting and appropriate instructional decisions and actions.

School Districts as Organizations

What are these special circumstances, people, or organizational arrangements? Within district evaluators or evaluation units would seem to have the potential for generating data that could be useful, for example, in measuring the effectiveness of the instructional program and recommending ways to improve teaching and learning. What are the obstacles to doing this?

One year ago, our CSE project, officially titled 'Evaluation Design: An Organizational Study," was funded by NIE to find districts that were purposefully using findings from externally mandated testing or evaluation efforts in a way that influenced their instructional decision making. Hence the informal title of our project--"Linking Testing and Evaluation with Instruction."

From the beginning, we knew from the literature and our own research and experience in school districts that district-forged links between testing, evaluation, and instruction are not commonplace. There may be

a number of reasons for this non-linkage related to factors such as the characteristics of mandated tests and evaluations, the role of evaluators, or technical problems with analysis. In addition, some factors within the school district's organizational setting may be influential. We will mention only four that seem plausible contributors to the non-use of test and evaluation data.

Loose coupling. This term refers to the degree to which the various units of any organization are coordinated with and dependent upon one another. For example, how likely is it that a decision made by top management will be implemented at the lower operational levels? Is the coordination among levels tight or loose? Typically within school districts the system linking board members and central administrators with classroom instructional activities is very loose. District level policy decisions relating to instruction may not be routinely implemented in the classroom (Goodlad & Klein, 1970; Meyer, 1977). The policy makers' intent may be misunderstood, changed, or ignored by classroom teachers. Thus, district administrators who want to increase teachers' classroom use of testing results, for example, will likely have to take extraordinary measures to effect such behavior.

Teacher isolation. Another reason why district policy level decisions related to instruction may not be carried out in the classroom is because teachers often work isolated behind closed doors from one another and from external supervision. Consequently, it is difficult for supervisors to influence the teacher's daily activities. (Lortie, 1975)

Permeable boundaries. School districts' organizational boundaries can and often are breached by external agencies --witness external regulations or mandates from the courts or the state and federal government.

Local interest groups can often put pressure on school district decision makers. Societal influences such as population shifts, increases in immigration, inflation, changing tax structures also affect districts. School districts, therefore, have to continually adjust their activities so as to meet changing and sometimes conflicting demands and priorities.

Goal ambiguity. In our pluralistic society undergoing rapid social change, the stated goals of schooling seem to shift with some rapidity. Not only do the goals change over time; they are often expressed in ambiguous terms as well. Unlike organizations in the private sector with a "profitability" bottom line, public schools must struggle with the difficulties of measuring their success "in developing responsible citizens," creating "safe drivers," or "students with an appreciation of their historical heritage." The students' attainment of these goals is difficult to chart. There is considerable disagreement in society about priorities and standards for educational achievement. Factors such as these appear to diminish the utility of test data as credible measures of a school district's success in educating children.

In view of these school district organizational characteristics, it seems unlikely that districts will naturally and easily integrate testing and evaluation data with instructional decision making. However, there are evaluators and administrators who believe that monitoring student learning and taking corrective instructional action where necessary is important.

Research Methods

Our project began looking about for districts that had such a reputation of linking evaluation or testing with instruction. Based on

previous CSE research, and conversations with colleagues in school districts, state departments of education, and other universities and research institutions, we identified several districts that were thought to have linked their testing and evaluation activities with instructional decision making in some manner. From 40 nominees, a final sample of four districts was selected. While these districts cannot be viewed as representative of all school districts, they do exhibit characteristics that represent the diversity of American school districts, e.g., differences in size (e.g., large-small), student demographics (e.g., affluent-below average income, racially homogeneous-racially heterogeneous) locale (urban-suburban).

For the past year we have conducted field work in these four districts. Our interviews and document analyses were built around three questions:

1. Why? If most districts are not trying to link testing and evaluation with instructional decision making, why are these districts the exception? What is the history of their efforts? Was it planned? Who were the critical actors, what were the critical events?
2. What? What has been the result? What kind of linking system do they have and how does it work? How mature or fully developed is the linkage between testing, evaluation, and instruction?
3. So what? Has there been any payoff from their efforts? Presumably their linkage system was developed to accomplish some purpose--what evidence is there that the linking system has had its intended effect?

At present we have completed the first year's work in four districts and are sifting through the data in an attempt to identify common properties and consistent patterns. While we have not yet completely answered our research questions, we have some preliminary observations we would like to share with you.

District's Linking of Testing and Evaluation with Instruction

Environmental context. It is self-evident that school districts exist in a social and historical context, as well as within a particular community. It is also self-evident, but sometimes overlooked, that the individuals working within school districts and classrooms are participants in the social and cultural ambiance of their times. They are also members of their professional educational communities, simultaneously shaping it and being shaped by it. What struck us forcibly about the four districts in which we worked was the influence these various environments had upon the school district personnel's thinking and actions.

One of the most striking similarities we noticed among our four districts was the amount of turmoil they were experiencing. One district was preoccupied with responding to court desegregation directives which necessitate district-wide management changes and changes in the autonomy teachers in minority isolated schools will have in instructional decision making. Another district was moving quickly to transform all its elementary schools into fundamental schools after a brief two-year period of pilot testing a fundamental school model in a few schools. This shift was in response to pressures from a conservative school board representing a community interested in a return to "basics." A third district was

struggling with a sudden increase in minority and non-English speaking students who added to an already diverse mix of students. The district was investing enormous time and energy in managing effective instruction for minority children with limited English-speaking capability. The fourth district, while not experiencing these particular societal tensions, was shifting from an elementary school district to a unified district, K-12. They were trying to incorporate secondary schools into the existing elementary school district structure. In each case, while inventing solutions to deal with these immediate problems, school officials felt no certainty that solutions or procedures invented for this year's problems would be appropriate for handling next year's problems.

In each of our four districts there is evidence of goal diffuseness and boundary permeability. The external environment has frequently invaded district boundaries--e.g., court mandates, demands for bilingual programs, population changes, unification elections--and forced the districts to redirect their energies. Under these circumstances, district goals have become diffused. District officials have difficulty in maintaining long-term consistency in ordering their priorities and pursuing their goals. Given these external conditions, district abilities to develop long-term plans would seem to be severely challenged.

In view of these factors, it would be surprising if testing of children for the purpose of evaluating and improving instruction was uppermost in the minds of school officials. While admittedly not their first concern, in the four districts we studied, testing and evaluation activities and their linkages to instructional improvement were receiving district-wide attention. Paradoxically, this seemed due to the same

pressures in the environment which were making planning difficult. For example, in the district moving toward fundamental schools, test scores were being considered by the board both as evidence of the effectiveness of the revised program and as a monitoring device for teachers' use in tracking student progress. In the small elementary district, a comprehensive criterion-referenced testing system had developed in response to community and administrative interests in individualizing instruction for students. In the heterogeneous district, State Assessment Program tests were being analyzed to see how the curriculum for various populations matched the specifications of the items. It seems that, in these districts, local environmental forces interacted with state and local requirements to influence particular individuals to take actions linking testing and evaluation with instruction.

Personnel. Another notable characteristic evident in the four districts we have studied is the professional interest key personnel have in instructional improvement.

Central office administrators in each of the four districts were a relatively stable group. In spite of changes at the superintendent level, the individuals responsible for curriculum, instruction, and supervision of elementary and secondary levels have worked together over a long period of time. In all four districts these individuals have evolved methods of communicating and resolving difficulties. This stability, rather than leading to stagnation, seems to have contributed in three of the four districts to a sense of direction more coherent than one would have thought possible given other organizational and environmental instabilities.

A testing/evaluation/instructional subsystem. Although the four districts differed from one another in their size, organization, and structure, we have sought to develop a conceptual framework to categorize some common aspects of their management of testing, evaluation, and instruction. Here we will only outline what we mean by such a management subsystem rather than provide district by district details.

We believe that the linking of testing and evaluation with instructional decision making can be characterized as a management subsystem--a testing, evaluation, and instruction (T/E/I) linking subsystem. Such a subsystem is not a formal structure that appears on the school district organizational chart; instead it is an alignment of individuals and groups that have, for a variety of reasons, made informal and formal arrangements that constitute this subsystem. The subsystem can consist of two people or a dozen people, depending on the size of the district and the way in which the subsystem is defined. It is not limited to those individuals necessarily concerned with testing and evaluation.

Three components seem to be necessary but not sufficient in order for this subsystem to form and operate: ideas, operations, and coordinating mechanisms.

Ideas. We mean those beliefs, goals, assumptions, sometimes acknowledged, sometimes not, etc., that guide the district's activities. In our four districts these were implicit and explicit ideas held by a few, some, or many people, within which the district evolved its subsystem. If the ideas were realistic, accurate, and complete, the subsystem had a greater likelihood of succeeding; if the ideas were faulty, incomplete, or incompatible with one another the subsystem faltered.

Perhaps a couple of examples will illustrate, in general, what we mean by "ideas." Teachers, as we have noted, work in relative isolation, beyond close supervision. They tend to accept and utilize those programs and projects that solve their immediate classroom teaching needs. If a district decides that they will mandate teacher attention to test scores in their classrooms and the district does not provide for reporting the test scores in a way that is of immediate practical use to teachers, then the district's ideas about its delivery system is faulty and it is almost certain that the teachers will not utilize the tests in the way the district intends.

Similarly, suppose a district board of education mandates that the CTBS test results are to be the dependant measure of a school district's educational program and the board does not assure that the district's instructional continuum is related to the content of the CTBS battery--then their ideas are faulty. The CTBS battery likely has some unknown relationship to the quality of the district's instructional program, but it is mischievious to try and link the two together. Any conclusions about program success or student progress will be invalid.

Operations. By this we mean those individuals, organizational arrangements, and technical capacities that enable the district to implement and sustain the district's ideas. Districts must have high quality and a wide range of operations in order to manage a T/E/I linking subsystem.

For example, a district may desire to provide test or evaluation results to teachers quickly so as to increase the results' practical value to classroom teachers. But if the district does not have the necessary computer programming skills or access to appropriate computer

facilities, then turn-around time remains a problem; the results cannot be reported in a useable time frame. Or if the district wishes teacher to take prescriptive action in relation to diagnostic testing, and if the teachers are not presently skilled in how to do that, then it is necessary to provide appropriate inservice. Such an inservice training program, then, by our definitions, would be an essential operation in linking testing and evaluation with instruction.

Coordinating mechanisms. As we have noted earlier, school districts have often been characterized as loosely coupled; that is, communications and coordination among the various subunits is often loose. In many districts, the curriculum division, and school principals, and the testing and evaluation unit are often surprisingly uninformed about each others' activities and problems. If a T/E/I linking subsystem is going to work it seems necessary that the various subunits and individuals who are going to be affected by the subsystem be brought together in some way for communications and/or decision-making purposes. In small districts this may be accomplished through somewhat informal means (although small size does not necessarily assure that it will occur); in large districts this may require establishing a cross-functional committee system. Whatever the specifics, some coordinating mechanisms among ideas and among operations seem essential to manage the linkages between testing, evaluation, and instruction.

Summary

We have presented some initial, relatively diffuse notions about the obstacles to, and the conditions needed for school districts' linking of externally mandated testing and evaluation with instruction. In

second and third year activities of the project, we will look to describe more completely the environmental and management factors which both impede and contribute to successful district utilization of data from tests and evaluations for locally initiated instructional improvement. The two major points, however, that we have tried to make in this paper are these:

- The evaluation and testing communities must attend to the needs of administrators and teachers as they are manifest in the district environment, if testing and evaluation efforts are to be used at the local level to improve instruction.
- The linking of testing and evaluation with instruction does not happen within districts without intention and effort. District-wide management subsystems composed of certain kinds of ideas. a range of related operations, and a variety of coordinating mechanisms seem to be needed. Questions relating to why such subsystems evolve, how they can operate effectively, and how they can be facilitated are worthy of continued attention and support.

References

Alkin, M. C., Daillak, R., White, P. Using evaluations: Does evaluation make a difference? Beverly Hills, Calif.: Sage Publications, 1979.

Callahan, R. C. Education and the cult of efficiency. Chicago: University of Chicago Press, 1962.

Daillak, R. H. Evaluators at work. Los Angeles: Center for the Study of Evaluation, UCLA Graduate School of Education, September 1980.

David, J. Local uses of Title I evaluation. Palo Alto, Calif.: Stanford Educational Policy Research Center, 1978.

Goodlad, J., & Klein, M. F. Behind the classroom door. Belmont, Calif.: Wadsworth Publishing Co., 1970.

Lyon, C. D., Doscher, L., McGranahan, & Williams, R. Evaluation and school districts. Los Angeles: Center for the Study of Evaluation, UCLA Graduate School of Education, 1978.

Meyer, J. W. Research on school and district organizations. Paper presented at annual Sociology of Education Association conference, San Diego, 1977.

Patton, M. Q. Utilization-focused evaluation. Beverly Hills, Calif.: Sage Publications, Inc., 1978.

Weiss, C. H. (Ed.). Evaluating action programs. Boston: Allyn & Bacon, 1972.